



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/804,516	03/19/2004	Arun Kwangil Iyengar	YOR920040025US1	7509

7590 12/03/2008
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560

EXAMINER

PHUNG, LUAT

ART UNIT	PAPER NUMBER
----------	--------------

2416

MAIL DATE	DELIVERY MODE
-----------	---------------

12/03/2008

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

DETAILED ACTION

Response to Amendment

1. Applicants' arguments filed on September 2, 2008 have been fully considered but they are not persuasive.

2. On page 3, applicants argue that:

It is clear that what Kokot discloses is that, as defined by a predefined subscriber profile, some particular packet flows may be treated preferentially and may have a QoS class associated therewith ("[t]he service profile may identify classes of packets that may be forwarded on preferential packet flows"). However, this does not mean that a client is assigned to a QoS class [emphasis original], as is expressly claimed.

Examiner disagrees because:

As stated in the rejection of claim 1, Kokot discloses a QoS class assigned to a subscriber from which the packet flow originated (**para. 116; a "service profile for a subscriber may include, for example, one or more general QoS classes for packet flows originating from ... a subscriber device associated with the subscriber"**). A service profile for a subscriber including a QoS class means the QoS class having been assigned to a subscriber. Furthermore, Kokot discloses "QoS profiles describe the QoS classes, if any, that subscribers ... are authorized to receive." (para. 104). Thus Kokot clearly anticipates the claimed limitation of "a client is assigned to a QoS class".

3. On page 4, applicants argue that:

In addition, Applicants, after considering the present Office Action in its entirety, respectfully assert the same deficiency arguments presented in their previous response

Art Unit: 2416

dated May 5, 2008 (the disclosure of which is incorporated by reference herein) with respect to Bender, Veres, Menditto and Lu. In view of the above, Applicants believe that claims 1-25 are in condition for allowance, and again respectfully request withdrawal of the various remaining rejections.

Examiner disagrees because:

The aforementioned arguments have been addressed in the last office action, to which applicants are requested to refer for responses.

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

Art Unit: 2416

6. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

7. Claims 1, 5-9 and 14-17 are rejected under U.S.C. 103(a) as being unpatentable over Colby, et al (US 6,449,647) in view of Kokot, et al (US 2004/0258003).

Regarding claim 1, Colby discloses an article of manufacture for processing a request to at least one server, comprising a computer readable medium containing one or more programs (**abstract; Fig. 23; col. 19, lines 47-65**) which when executed implement the steps of:

receiving a request (**col. 2, lines 55-56**); and

scheduling submission of the request to the at least one server (**col. 2, lines 58-59**) based on: (i) a quality-of-service (QoS) class (**col. 9, lines 36-39**); (ii) a response target associated with the QoS class (**col. 9, lines 25-32; col. 21, lines 50-51**); and (iii) an estimated response time associated with the at least one server (**col. 22, lines 10-11**).

Colby discloses all of the subject matter except (i) a quality-of-service (QoS) class assigned to a client from which the request originated. Kokot from the same or similar fields of endeavor discloses a QoS class assigned to a subscriber from which the packet flow originated (**para. 116; a “service profile for a subscriber may**

Art Unit: 2416

include, for example, one or more general QoS classes for packet flows originating from ... a subscriber device associated with the subscriber”). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine Colby's method of processing a request based on QoS flow classes with Kokot's QoS subscriber classes by assigning the QoS classes to the clients and scheduling the request based on these QoS classes. The motivation for doing so would have been to ensure proper level of quality for the service request.

Regarding claim 5, Colby in view of Kokot further discloses further comprising the step of assigning the response target to the QoS class **(Table 1; col. 9, lines 25-39; col. 22, lines 10-11)**.

Regarding claim 6, Colby in view of Kokot further discloses wherein the step of assigning the response target to the QoS class further comprises the step of assigning a response time target to the QoS class. **(Table 1; col. 9, lines 25-39; col. 22, lines 10-11)**

Regarding claim 7, Colby in view of Kokot further discloses wherein the step of assigning the response target to the QoS class further comprises the step of assigning a response percentile target to the QoS class. **(col. 16, lines 56-61 and 65-66)**

Regarding claim 8, Colby in view of Kokot further discloses further comprising the step of estimating the response time associated with the at least one server based on one or more requests sent to the at least one server within a given time period. **(col. 6, line 46 to last line; claim 32)**

Regarding claim 9, Colby in view of Kokot further discloses QoS classes based on delay, whereas lower quality class is assigned higher delay (**QoS class 3 with delay of 500 ms per Table 1**) than higher quality class (**QoS class 1 with delay of <250 ms per Table 1**). Examiner takes official notice that it is well known in the art that delay is proportional to response time in terms of processing a request by a server.

Regarding claim 14, Colby further discloses an apparatus for processing a request to at least one server, comprising:

a memory (**RAM per Fig. 23, element 110**); and

at least one processor coupled to the memory (**Fig. 23, element 1080**) and operative to perform the method of claim 1, and is therefore rejected under the same reason set forth in the rejection of claim 1.

Regarding claim 15, Colby in view of Kokot further discloses wherein the memory and the at least one processor form a scheduler (**Fig. 1b, element 110**) that is external to the at least one server (**Fig. 1b, elements 100**).

Regarding claim 16, Colby in view of Kokot further discloses the apparatus of claim 15, wherein the scheduler is a front-end scheduler and the at least one server is a back-end server (**Fig. 1b, elements 110 and 100; col. 16, lines 28-29**).

8. Claims 2-4, 18-20 and 25 are rejected under U.S.C. 103(a) as being unpatentable over Colby, et al in view of Kokot, et al, and further in view of Bender, et al (US 6,112,221).

Regarding claim 2, the combination of Colby and Kokot discloses all of the subject matter as disclosed previously in this office action except for the following:

further comprising the step of withholding the request from submission to the at least one server when the request originated from a client assigned to a first QoS class to allow a request that originated from a client assigned to a second QoS class to meet a response target associated therewith.

Bender from the same or similar fields of endeavor discloses a server which employs a pre-emptive setting not continuously processing a request, but scheduling them according to an earliest deadline first methodology, by alternately processing the request with the earliest deadline first, followed by that with the next earliest deadline, and so on (**col. 4, lines 52-58; col. 5, lines 27-35**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the pre-emptive scheduling method of Bender by pre-empting the request having a lower QoS class with the request having a higher QoS class. The motivation for using pre-emptive procedure would have been to prioritize requests according to response target.

Regarding claims 3 and 4, the combination of Colby, Kokot and Bender discloses substantially all of the subject matter as disclosed previously in this office action. Colby further discloses:

determining a throughput of the at least one server (**col. 6, line 46 to last line**), as recited in claim 3; and

monitoring a throughput of the at least one server (**col. 6, line 46 to last line**), as recited in claim 4.

Colby and Kokot do not explicitly disclose:

reducing a request withhold rate to increase throughput of the at least one server, as recited in claim 3;

varying a request withhold rate to balance the throughput and request response times, as recited in claim 4.

Bender from the same or similar fields of endeavor discloses a server which employs a pre-emptive setting of scheduling requests according to an earliest deadline first methodology (**col. 4, lines 52-58; col. 5, lines 27-35**), calculating processing time and dead line for each request (**Fig. 2, element 102**), and continue adjusting estimated processing time (**Fig. 2, element 112**).

Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the pre-emptive scheduling method of Bender by pre-empting the request having a lower QoS class with the request having a higher QoS class, monitoring the throughput of the servers, and adjusting the pre-emption rate so that the request can be processed within the target response time. The motivation for using pre-emptive procedure would have been to prioritize requests according to response target.

Regarding claims 18 and 25, Colby discloses assigning a response target to at least one QoS class (**Table 1; col. 9, lines 25-39; col. 22, lines 10-11**); and estimating at least one response time of the at least one server based on one or more requests

Art Unit: 2416

sent to the server within a given time period (**col. 2, lines 54-64; col. 6, lines 46 to last line; claim 32**)

Colby does not explicitly disclose assigning at least one client to a quality-of-service (QoS) class from among at least two QoS classes. Kokot from the same or similar fields of endeavor discloses assigning at least one client to a quality-of-service (QoS) class from among at least two QoS classes (**para. 116**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine Colby's method of processing a request with Kokot's QoS subscriber classes by assigning the QoS classes to the clients. The motivation for doing so would have been to ensure proper level of quality for the service request.

The combination of Colby and Kokot discloses all of the subject matter except: withholding requests associated with a first one of the at least two QoS classes to allow requests associated with a second one of the at least two QoS classes to meet its response target based on the at least one estimated response time.

Bender from the same or similar fields of endeavor discloses a server which employs a pre-emptive setting not continuously processing a request, but scheduling them according to an earliest deadline first methodology, by alternately processing the request with the earliest deadline first, followed by that with the next earliest deadline, and so on (**col. 4, lines 52-58; col. 5, lines 27-35**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the pre-emptive scheduling method of Bender by pre-empting the request having a lower QoS class with the request having a

Art Unit: 2416

higher QoS class. The motivation for using pre-emptive procedure would have been to prioritize requests according to response target.

Claims 19 and 20 are substantial duplicates of claims 3 and 4, respectively, and are therefore rejected under the same reason set forth in the rejection of claims 3 and 4, respectively.

9. Claim 8 is rejected, in an alternative, under U.S.C. 103(a) as being unpatentable over Colby, et al in view of Kokot, et al, and further in view of Veres, et al (US 6,807,156).

Regarding claim 8, the combination of Colby and Kokot discloses all of the subject matter as previously recited in this office action. Further, Veres from the same or similar fields of endeavor discloses further comprising the step of estimating the response time (**col. 13, lines 46-47**) associated with the at least one server or applications based on one or more requests sent to the at least one server or applications within a given time period (**time window of measurement as shown in Fig. 2; col. 13, lines 36-47**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the response time estimating method of Veres by periodically sending requests to the applications and servers to estimate the response time. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

Art Unit: 2416

10. Claims 10-12 are rejected under U.S.C. 103(a) as being unpatentable over Colby, et al in view of Kokot, et al, and further in view of Menditto, et al (US 6,981,029).

Regarding claim 10, the combination of Colby and Kokot discloses all of the subject matter as disclosed previously in this office action except for the following:

determining dispatch times for requests from a difference between at least one predicted response time of the at least one server and the target response time corresponding to the QoS class of the request; and

sending requests to the at least one server based on dispatch times.

However Colby discloses using response time as a metric to select a server to meet QoS requirements (**col. 9, line 36 to col. 10, line 59; claim 32**). Menditto from the same or similar fields of endeavor discloses a content gateway making routing decisions based on the request, selecting a server satisfying the request and depending on various factors such as server load, and forwarding the request to the selected server (**col. 3, lines 11-61**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

Regarding claim 11, Colby in view of Kokot further discloses:

wherein a plurality of applications are running on the at least one server and requests are routed to applications (**col. 1, line 55 to col. 2, line 4**), further comprising the steps of:

Art Unit: 2416

estimating response times of applications based on one or more requests sent to the applications within a time period. **(col. 6, line 46 to last line; claim 32)**

The combination of Colby and Kokot does not explicitly disclose:

sending a request to an application whose estimated response time is not greater than a target response time corresponding to the QoS class of the request.

However Colby discloses using response time as a metric to select a server to meet QoS requirements **(col. 9, line 36 to col. 10, line 59; claim 32)**. Menditto from the same or similar fields of endeavor discloses selecting an optimal server based on a set of rules, defining as producing the quickest response time to the request. **(col. 6, lines 16-40)** Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

Regarding claim 12, the combination of Colby and Kokot discloses all of the subject matter as recited above except:

further comprising the step of varying a number of requests sent to applications so that estimated response times of applications are not greater than target response times of QoS classes corresponding to requests sent to the applications.

However Colby discloses sending requests to applications to a server based on response time **(col. 6, line 46 to last line; claim 32)** Menditto from the same or similar fields of endeavor discloses a content gateway updating policies regarding processing

Art Unit: 2416

of requests based on service level agreements (**col. 7, lines 1-52**). Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the request processing method of Colby and Kokot with the QoS enforcement approach by Menditto by selecting a server that can timely process the request. The motivation for such a combination would have been to ensure service level agreement based on response time is met.

11. Claim 13 is rejected under 35 U.S.C. 103(a) as being unpatentable over Colby, et al in view of Kokot, et al, and Menditto, et al, and further in view of Lu, et al (US 6,772,211).

Regarding claim 13, the combination of Colby, Kokot and Menditto discloses all of the subject matter as disclosed previously in this office action except wherein the at least one server comprises a plurality of servers and each application runs on a different one of the plurality of servers.

Lu from the same or similar fields of endeavor discloses methods to switch client packets to one server among a group of servers (**col. 4, lines 50-53**) and applications have their own dedicated servers (**col. 5, lines 24-26**).

Thus it would have been obvious to the person of ordinary skill in the art at the time of the invention to combine the packet processing method of Colby, Kokot and Menditto with the servers and applications of Lu by implementing the method and each application on a separate server. The motivation for such a combination would have been to modularize the features for scalability and performance.

12. Claims 21-23 are rejected under U.S.C. 103(a) as being unpatentable over Colby, et al and Kokot, et al in view of Bender, et al and Menditto et al.

Claims 21-23 are substantial duplicates of claims 10-12, respectively, and are therefore rejected under the same reason set forth in the rejection of claims 10-12, respectively.

13. Claim 24 is rejected under U.S.C. 103(a) as being unpatentable over Colby, et al, and Kokot, et al in view of Bender, et al and Menditto et al in further view of Lu, et al.

Claim 24 is a substantial duplicate of claims 13 and is therefore rejected under the same reason set forth in the rejection of claim 13.

Conclusion

14. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Art Unit: 2416

15. Any inquiry concerning this communication or earlier communications from the examiner should be directed to LUAT PHUNG whose telephone number is (571) 270-3126. The examiner can normally be reached on M-Th 7:30 AM - 5:00 PM, F 7:30 AM - 4:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Ricky Q. Ngo can be reached on 571-272-3139. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/L. P./

Examiner, Art Unit 2416

/Ricky Ngo/

Supervisory Patent Examiner, Art Unit 2416